



# ProactiveAgent: Personalized Context-Aware Reminder System

Yumeng Ma\*

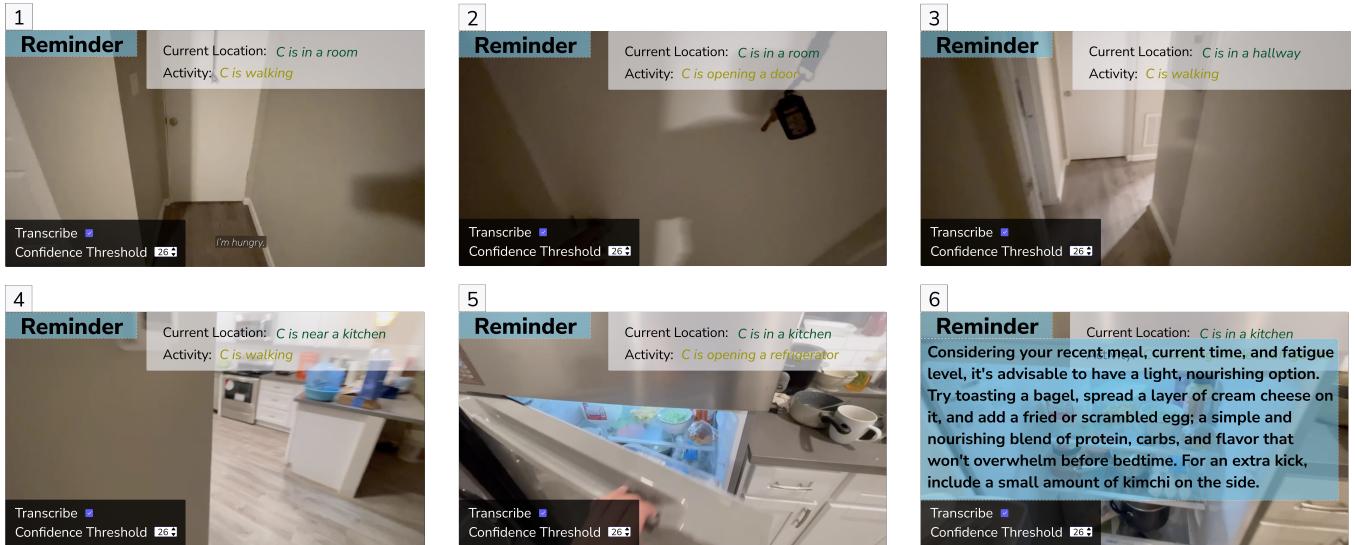
Jiahao Ren\*

yumeng\_ma1@brown.edu

jiahao\_ren@brown.edu

Brown University

Providence, Rhode Island, USA



**Figure 1:** Sequence of ProactiveAgent’s video stream frames depicting a user scenario: User C opening the refrigerator at night. By taking into account C’s recent meal (two hours ago), day of work, and the current time (late evening), the system proactively prompts C with a light snack suggestion from the items in front.

## ABSTRACT

We introduce ProactiveAgent, a proactive application that harnesses the capabilities of large language models (LLMs) and personal agents to provide context-aware, personalized reminders and suggestions. By assimilating real-time environmental data, user histories, and verbal interactions, the system discerns user intent and offers tailored recommendations. The application captures visual activity and spoken interactions, integrating them into short and long-term memory storage for context-rich decision support. We propose scenarios where ProactiveAgent could be valuable: suggesting snack options depending on the time, offering culinary options based on dietary preferences, and even guiding users in their daily tasks. In envisioned use cases, ProactiveAgent could potentially track user attributes during their shopping experience, such as time

\*Both authors contributed equally to this research.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*UIST ’23 Adjunct, October 29–November 01, 2023, San Francisco, CA, USA*

© 2023 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0096-5/23/10.

<https://doi.org/10.1145/3586182.3625115>

spent on items and other cues, leading towards insightful product recommendations. Our work represents a potential advancement in the realm of personalized assistance, merging LLM strengths with personal agent technologies to enhance user decision-making in dynamic real-world scenarios.

## CCS CONCEPTS

- Human-centered computing → Interactive systems and tools.

## KEYWORDS

Human-AI interaction, generative AI, large language models

## ACM Reference Format:

Yumeng Ma and Jiahao Ren. 2023. ProactiveAgent: Personalized Context-Aware Reminder System. In *The 36th Annual ACM Symposium on User Interface Software and Technology (UIST ’23 Adjunct), October 29–November 01, 2023, San Francisco, CA, USA*. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3586182.3625115>

## 1 INTRODUCTION

In the fast-paced landscape of technology, the demand for personalized user assistance has become ubiquitous, propelled by the emergence of generative AI and intelligent assistants such as Amazon

Echo and Google's Virtual Assistant. These advancements highlight the desire for tailored user support, encouraging the development of technologies to enhance the decision-making processes.

Recent research has demonstrated the potential of combining large language models (LLMs) with generative agents to simulate human behaviors, including planning and natural conversation [2]. Building upon this foundation, our work translates these findings into a practical real-world application, focusing on guiding human users toward logical and productive decisions. Our application also draws inspiration from InstructBlip [1]. Their work addresses the feasibility of abstracting visual information into textual data, enabling users to interact with their surroundings using language. Our application extends this concept by employing LLMs to facilitate user interaction with the environment. LLMs have demonstrated their capacity for maintaining long-term memory and rephrasing text into instructional language. By utilizing these properties, we can have a more informative and logical assistant to teach users how to cook and use new tools.

We introduce ProactiveAgent, a proactive reminder application that leverages advancements in LLMs and personal agents to enhance user decision-making processes in their daily lives. The application's ability to assimilate real-world surroundings, past interactions, and speech provides suggestions that reflect a user's unique context and preferences.

To illustrate how the proactive reminder system works, the user wears smart glasses to allow the collection of their visual activity data throughout the day. The integration of speech-to-text transcription also collects any spoken interactions. All the information are stored for reference.

For instance, consider a scenario where the system has insight into the user's recent actions—they ate two hours ago, are currently hungry, and have been engaged in work all day. At present, the user indicates their hunger while walking from their room to the kitchen. In response, the system recommends a light snack option that allows minimal preparation, addressing their prompt needs (Figure 1).

In a dietary guidance context, users present their food preferences, allergies, or special occasions like birthdays. The system, using short-term memory inputs, tailors recommendations to meet these specific criteria while ensuring the user's well-being.

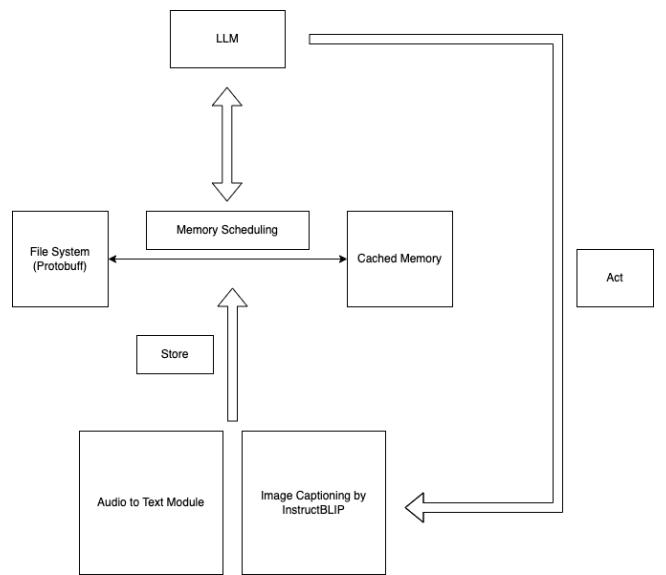
Visualize another scenario where the application senses that the user is at their favorite coffee shop and recalls that the user typically orders a latte on Monday mornings. Not only does it remind the user of this preference, but it might also suggest trying a new flavor based on a recent article they read about coffee. This synthesis of user behavior insights, combined with dynamic experiences, ensures that ProactiveAgent's suggestions are both relevant and timely.

To showcase ProactiveAgent's ability, we propose a demonstration where we assimilate predefined data from a live audience into the system's short-term memory. Subsequently, users are immersed in a simulated shopping experience. As the activity unfolds, ProactiveAgent tracks attributes like time allocation on specific items and both verbal and non-verbal indicators of interest. Synthesizing this data, the system generates product recommendations resonating most with the user's observed inclinations.

## 2 SYSTEM DESIGN

### 2.1 Pipeline

The pipeline of the system consists of a two-fold approach: memory scheduling and suggestion proposal (Figure 2). Memory scheduling is a hierarchical structure that comprises of both long-term and short-term memory. Upon receiving input data, encompassing video frames and audio transcriptions, the system decides whether this information is stored in the long-term memory or is better suited for the short-term memory. The summary module continually refines the stored data, trimming redundancy and outdated content. The suggestion proposal dynamically generates suggestions and responds to user intent, leveraging input data with the insights drawn from the combination of long-term and short-term memory banks.



**Figure 2: ProactiveAgent's pipeline.**

### 2.2 User Interface

The interface of the application includes the following:

- *Camera Preview* displays what the user currently sees.
- *Reminder Component* displays proactive reminders based on user contextual cues.
- *Current Location and Activity Elements* showcases the user's current location and ongoing activity.
- *Parameter Configuration* provides the user with the ability to adjust the application's parameters, enabling actions like toggling transcription and setting confidence thresholds.
- *Real-Time Transcription* of user speech can be displayed in real time.

## ACKNOWLEDGMENTS

We thank Jiahao "Nick" Li and Xiang 'Anthony' Chen from the UCLA HCI Lab for their support and guidance.

## REFERENCES

- [1] Wenliang Dai, Junnan Li, Dongxu Li, Anthony Meng Huat Tiong, Junqi Zhao, Weisheng Wang, Boyang Li, Pascale Fung, and Steven Hoi. 2023. InstructBLIP: Towards General-purpose Vision-Language Models with Instruction Tuning. In *Proceedings of the ACM International Conference on Multimedia (ACM MM '23)*. arXiv:2305.06500 [cs.CV]
- [2] Joon Sung Park, Joseph C O'Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. *arXiv preprint arXiv:2304.03442* (2023).

## A ETHICAL STATEMENT

We acknowledge that the proactive reminder application utilizes personal data and accesses sensitive information like user history

and real-time context. The algorithms driving personalized recommendations might inadvertently reinforce biases from training data, leading to unfair suggestions. We are also aware that our application's proactive nature could influence decisions without explicit consent. Societal impacts include potential over-reliance on its suggestions and diminishing critical thinking skills. To address these, we commit to implementing robust privacy measures that involve regular security audits and ensuring that long-term memory is always stored locally. We will continue to fine tune our application and employ debiasing techniques, ongoing evaluation, and corrective actions to ensure fairness. User well-being will be central, allowing customization, feedback, resets, and overrides.